Efficient Calculation of Certain Integrals

For Modelling Extremely Rare Events

Peter Robinson

Quintessa Limited.
The Hub, 14 Station Road,
Henley on Thames,
Oxfordshire,
UK RG9 1AY
PeterRobinson@quintessa.org

September 2009

Beauzamy (2009) proposes a novel approach to modelling rare events. The details are not important here; we focus on the evaluation of two integrals, or more precisely of their ratio. In the more difficult example given by Beauzamy, the integrals are 32 dimensional and take many hours to evaluate using an algebraic approach.

We present a method that calculates the same integrals in just a few seconds on a standard desktop computer using a numerical approximation. Moreover our method will scale well for larger systems.

The two integrals of interest are

$$J = \int_S x_1^{n_1} \ldots x_k^{n_k} \ldots x_K^{n_K}\, dx_1 \ldots dx_K \tag{1}$$

$$J_k = \int_S x_1^{n_1} \ldots x_{k-1}^{n_{k-1}} x_k^{n_k+1} x_{k+1}^{n_{k+1}} \ldots x_K^{n_K}\, dx_1 \ldots dx_K \tag{2}$$

where S is the simplex given by

$$x_1 \geq x_2 \ldots \geq x_K \geq 0, \;\; \sum_k x_k = 1 . \tag{3}$$

The required results are

$$\overline{p}_k = \frac{J_k}{J} . \tag{4}$$

We use a stochastic approach. We generate points in the simplex S at a constant density and then approximate the required results as

$$\overline{p}_k = \lim_{N\to\infty} \frac{\sum_{i=1}^{N} x_{1,i}^{n_1} \dots x_{k-1,i}^{n_{k-1}} x_{k,i}^{n_k+1} x_{k+1,i}^{n_{k+1}} \dots x_{K,i}^{n_K}}{\sum_{i=1}^{N} x_{1,i}^{n_1} \dots x_{k,i}^{n_k} \dots x_{K,i}^{n_K}} . \tag{5}$$

To generate the points in S at constant density, we proceed as follows. For each point

i. Generate K-1 uniformly distributed random numbers in [0,1);

ii. Order these and call them $U_1, U_2, \dots U_{K-1}$.

iii. Calculate the spacings $d_1, d_2, \dots d_K$ where $d_1=U_1$, $d_k=U_k-U_{k-1}$, $d_K=1-U_{K-1}$.

iv. Sort these in descending order, and they are the required x components.

Devroye (1981) states that it is well-known that the spacings are uniformly distributed in the required simplex.

All the terms in the sums in (5) are positive so cancellation is not a problem. The number of samples required for convergence has been experimentally found to be a few million. The evaluation is cheap, so an over-generous number of samples can be used.

The example used is the harder case from Beauzamy (2009). The values for the n_k are taken from Table 2 of that paper (noting that the last zero value is omitted there).

n_1	9	n_{12}	4	n_{23}	3
n_2	7	n_{13}	1	n_{24}	0
n_3	8	n_{14}	2	n_{25}	0
n_4	11	n_{15}	2	n_{26}	0
n_5	6	n_{16}	1	n_{27}	0
n_6	3	n_{17}	1	n_{28}	1
n_7	7	n_{18}	1	n_{29}	0
n_8	6	n_{19}	2	n_{30}	0
n_9	6	n_{20}	0	n_{31}	0
n_{10}	4	n_{21}	0	n_{32}	0
n_{11}	3	n_{22}	0		

The results are given below for N=5 000 000. Five decimal places are given. All results agree with the 4 decimal places given by Beauzamy (Table 4). This took 15 seconds to calculate on a standard desktop computer.

1	0.11697	12	0.03269	23	0.00930
2	0.09653	13	0.02870	24	0.00793
3	0.08608	14	0.02567	25	0.00675
4	0.07797	15	0.02297	26	0.00571
5	0.06855	16	0.02048	27	0.00478
6	0.06122	17	0.01836	28	0.00394
7	0.05629	18	0.01649	29	0.00300
8	0.05130	19	0.01479	30	0.00215
9	0.04639	20	0.01306	31	0.00137
10	0.04127	21	0.01162	32	0.00066
11	0.03665	22	0.01038		

We note that the scheme requires a good quality random number generation. We used the "SIMD-oriented Fast Mersenne Twister" (SFMT) from http://www.agner.org/random/.

We also note that the scheme will handle much larger cases since the computation time is O(Kln(K)). We note that it may be necessary to scale the terms in the sums in (5) to avoid underflows; this was not necessary in the example.

References

Beauzamy, B (2009). Robust Mathematical Models for Extremely Rare Events. Société de Calcul Mathématique, S.A. Algorithmes et Optimisation.

Devroye, L (1981). Laws of the iterated logarithm for order statistics of uniform spacings. The Annals of Probability Vol. 9, No. 5, 860-867.